



Détection Aveugle de Pulses Géants: Implantation sur GPU

Blind Detection of Giant Pulses: GPU Implementation

D. Ait-Allal, R. Weber***, C. Dumez-Viou*, I. Cognard***, G. Theureau*****

**Observatoire de Paris – Station de radioastronomie de Nançay, F-18330 Nançay, France
{dalal.ait_allal, cedric.dumez-viou, gilles.theureau}@obs-nancay.fr*

***Laboratoire PRISME, Université d'Orléans, Site Galilée, 12 rue de Blois, 45067 Orléans cedex 2, France
rodolphe.weber@univ-orleans.fr*

****Laboratoire de Physique et Chimie de l'Environnement et de l'Espace, UMR 6115 CNRS
F-45071 Orléans Cedex 02, France, icognard@cnrs-orleans.fr*

Mots clés: Élimination des RFI, Pulsar, détection de signaux impulsionnels, GPU.

Key words: RFI mitigation, Pulsar, transient detection, GPU.

Résumé

L'observation radio des Pulsars nécessite une instrumentation spécifique et des procédures de traitement du signal dédiées qui corrigent les effets de la dispersion induite par le milieu interstellaire. En outre, la qualité des observations peut être notablement dégradée par la présence d'interférences radioélectriques (RFI) d'origine anthropique. Ce papier présente l'instrumentation mise en place pour l'observation des pulsars à la Station de Radioastronomie de Nançay. Notamment, nous détaillerons une approche originale pour la détection automatique de pulses géants. Bien que moins sensible que l'approche classique consistant à balayer en temps différé l'espace des paramètres du pulsar potentiel, l'approche proposée se distingue par une efficacité d'implantation, une capacité de traitement en temps réel et une robustesse intrinsèque aux RFI. En outre, elle ne nécessite aucune connaissance préalable des paramètres du pulsar, ce qui autorise son utilisation pour la détection d'événements impulsionnels non répertoriés. Une implantation sur GPU est en cours de validation.

Introduction

Les pulsars sont des étoiles à neutron en rotation très rapide (jusqu'à 716 tours par seconde) qui émettent un faisceau radio balayant l'espace comme le faisceau d'un phare côtier. Si le faisceau est orienté en direction de la Terre, des pulses périodiques vont pouvoir être détectés par des radiotélescopes [1]. Cependant, lors de leur trajet dans l'espace interstellaire, ces pulses sont progressivement atténués et étalés temporellement (i.e. Les fréquences hautes vont arriver plus tôt que les fréquences basses). Aux figures 1.a et 1.b, la représentation temps-fréquence d'un de ces pulses est donnée: le pulse est difficilement détectable du fait de la dispersion. L'observation des pulsars requiert donc une instrumentation très spécifique, avec une haute résolution à la fois en fréquence et en temps, permettant de compenser la dispersion subie par le pulse. Les figures 1.c et 1.d illustrent l'effet positif de la dédispersion sur la détectabilité du pulse. De plus, la qualité des observations peut être dégradée par la présence (croissante) d'interférence radio d'origine humaine (RFI). Des algorithmes de détection de RFI peuvent être alors mise en œuvre pour éliminer les données corrompues. Quelques exemples de ces algorithmes sont donnés dans [2,3].

A la section 1, le récepteur à dédispersion cohérente développé à la station de Radioastronomie de Nançay est présenté. Il permet l'observation de pulsar connu dont on a déjà évalué le facteur de dispersion (DM, contenu intégré de la densité électronique le long de la ligne de visée) et la période.

La recherche de nouveau pulsar est une tâche délicate étant donné que ni la mesure de le DM, ni la périodicité du

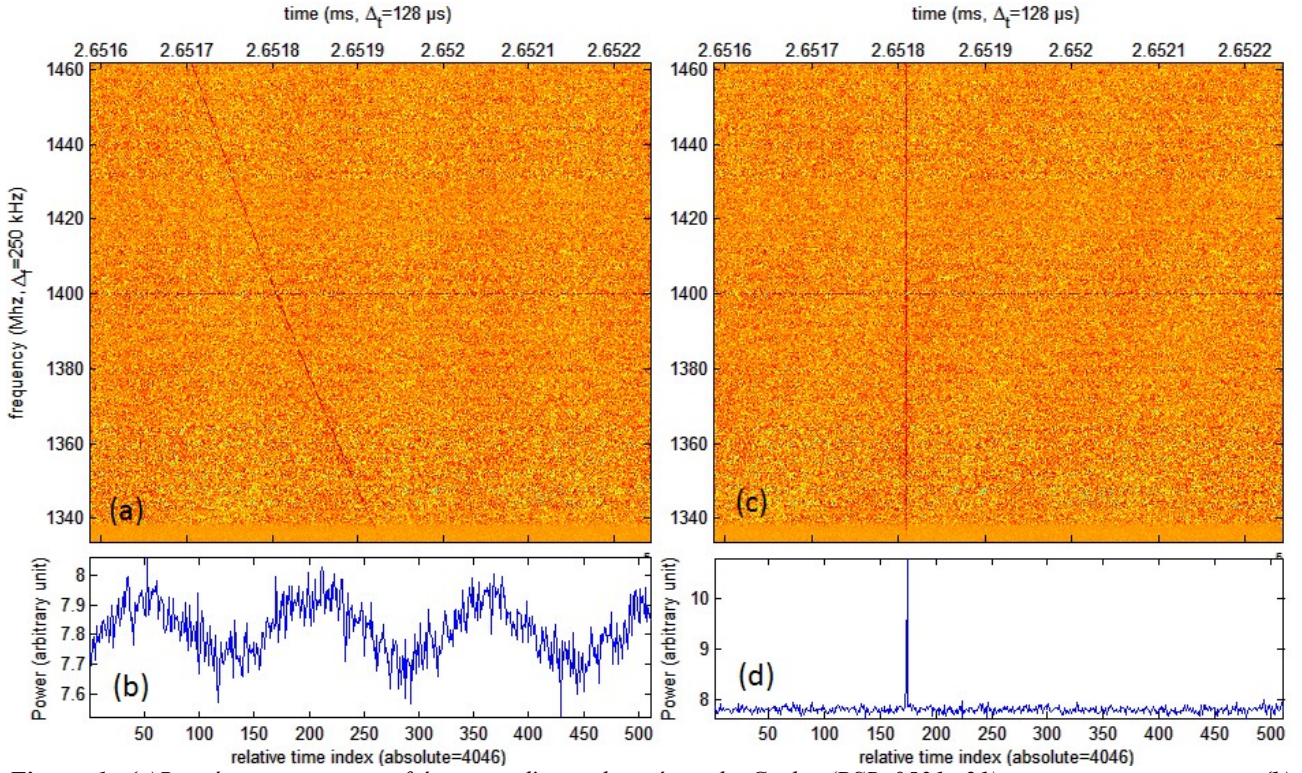


Figure 1: (a) Représentation temps-fréquence d'un pulse géant du Crabe (PSR 0531+21) avant tout traitement. (b) Puissance instantanée correspondante sur la bande totale. le pulse est indétectable. L'ondulation de la puissance instantanée est due au 50Hz secteur (c) Représentation temps-fréquence après dédispersion. (d) Puissance instantanée correspondante sur la bande totale. Le pulse est facilement détectable.

pulsar ne sont connues. L'approche classique consiste à enregistrer les signaux temporels puis à tester itérativement un ensemble de DM et de périodes. Évidemment, cette approche nécessite énormément d'espace de stockage puis de temps de calcul. En outre, elle ne dispose pas de robustesse intrinsèque aux RFI.

A la section 2, nous proposons une méthode originale de détection en aveugle (i.e. sans connaissance du DM ni de la période) et en temps réel de pulses. Elle ne nécessite moins d'espace de stockage et de ressource de calcul que l'approche classique. Par ailleurs, elle possède des capacités intrinsèques de robustesse face aux RFI. La contrepartie est une sensibilité moindre par rapport à l'approche classique. Elle est donc particulièrement bien adaptée à la détection de pulsars aux caractéristiques extrêmes dont l'existence a été découverte depuis quelques années: les RRATS [4] dont on ne reçoit qu'une impulsion de temps en temps (quelques par minute ou heure) et les «pulses géants» qui se caractérisent par des impulsions à l'intensité des dizaines à des milliers de fois plus intense que la moyenne.

Afin de démontrer son intérêt opérationnel, cette nouvelle technique est en cours d'implantation sur des GPUs. L'architecture de cette implantation est décrite à la section 3.

1. Le récepteur à dédispersion cohérente

1.1 Le principe de de-dispersion cohérente

La dispersion produite par la composante ionisée du milieu interstellaire peut être modélisée par une fonction de transfert [1], H :

$$H(f_0 + f) = \exp\left(+j \frac{2\pi DM f^2}{2.41 \times 10^{-10} (f + f_0) f_0^2}\right)$$

où DM est la mesure de dispersion, f_0 est la fréquence centrale de la bande Δf observée avec $|f| < \Delta f/2$.

La dédispersion cohérente consiste à appliquer la fonction de transfert inverse au signal reçu. Ce traitement est réalisé dans le domaine fréquentiel par le biais de transformées de Fourier. Le détail est donné dans le cadre intitulé «Data server» de la figure 2.

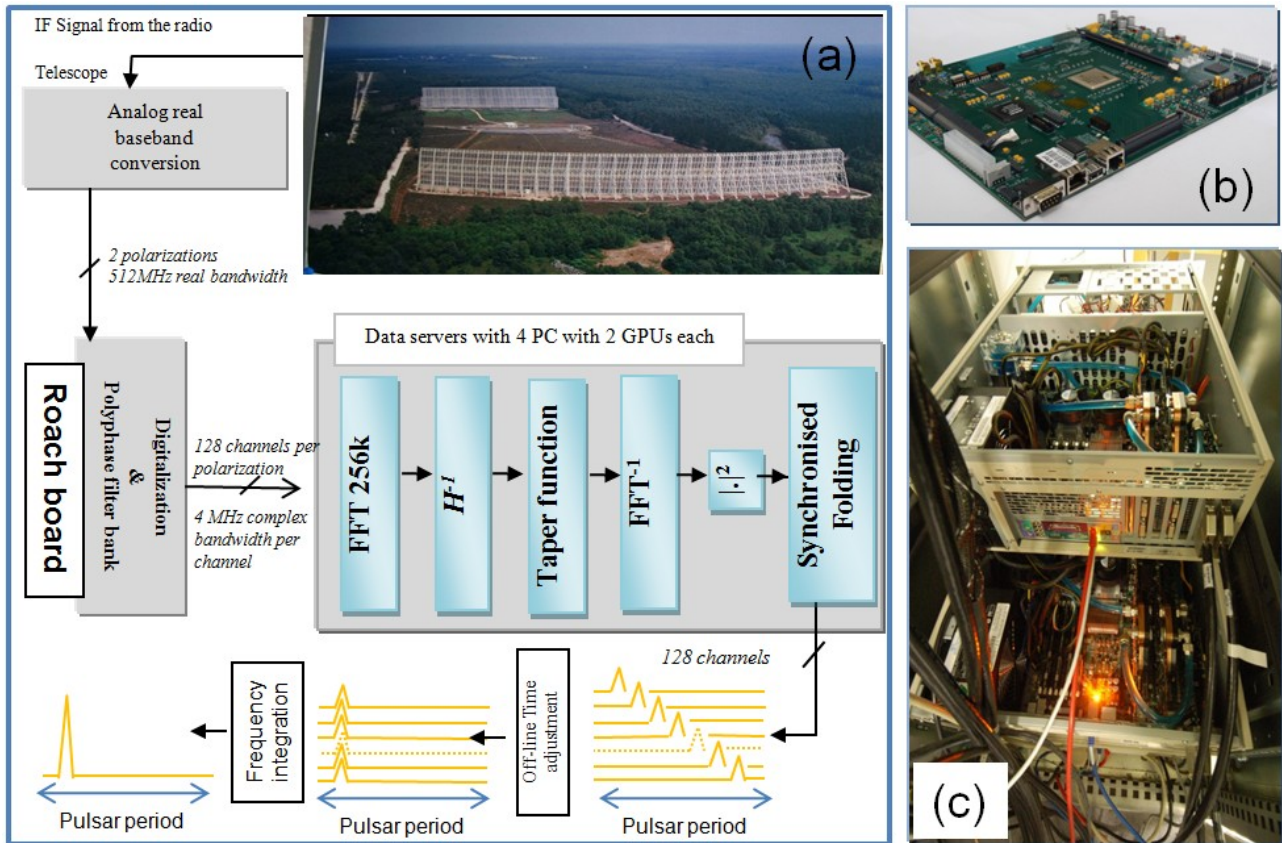


Figure 2: (a) Description de l'instrumentation pour l'observation des pulsars mise en œuvre à la Station de radioastronomie de Nançay sur son radiotélescope en bande décimétrique. (b) Carte numérique ROACH pour l'acquisition et le banc de filtres polyphases (c) « Data servers » avec les 4x2GPUs qui réalisent la dédispersion cohérente et le moyennage synchronisé des pulses (folding).

1.2 Implémentation matérielle

La figure 2.a montre l'ensemble du traitement à partir du radiotélescope jusqu'aux pulses dédispersés. Les différentes étapes sont décrites ci-dessous:

- la partie analogique: 2 bandes de 400MHz correspondant aux 2 polarisations orthogonales du radiotélescope sont descendues en fréquences afin d'être échantillonnées.
- la carte Roach (cf. figure 2.b et [6]): Cette carte numérique a été développée par CASPER (Center For Astronomy Signal Processing and Electronics Research). Elle numérise sur 8 bits les 2 signaux à 1024Mec/s et elle applique un banc de filtres numériques à 128 canaux. La largeur de bande des canaux de sortie est de 4 MHz. Les sorties complexes de chaque canal pour chaque polarisation sont codées sur 2 x 8 bits.
- Le serveur de donnée (cf. figure 2.c): 4 PC équipés de 2 GPU (GT280) récupèrent chacun un quart des différents canaux, soit 128 MHz, via des liens 10Gb et appliquent sur chacun d'eux la fonction de dédispersion appropriée via des Transformées de Fourier sur 256k points. Par défaut, les profils de pulses sont moyennés de manière synchrone sur 60s.

Cette configuration matérielle permet d'améliorer la sensibilité du récepteur d'un facteur 4 par rapport à la configuration précédente présentée dans [2].

2. Description du détecteur aveugle de pulses géants

A la figure 3.a, les différentes étapes de l'algorithme sont présentées et aux figures 3.b, 3.c et 3.d un exemple illustré sur le pulse réel de la figure 1.a est fourni. Pour résumer:

1. Un banc de filtres fournit un spectre temps-fréquence de dimension $N \times N$ (N fréquences et N instants) comme celui donné en exemple à la figure 1.a avec $N=512$.
2. Une FFT à 2 dimensions (FFT 2D) est appliquée sur ce plan temps-fréquence. Par ce biais, les pulses du pulsar sont tous projetés au même endroit en sortie de la FFT2D. Cet endroit, qui dépend de la dispersion subie par le pulse, sera différent de celui correspondant aux RFI. En effet, tous les RFI qu'ils soient à large bande ou à

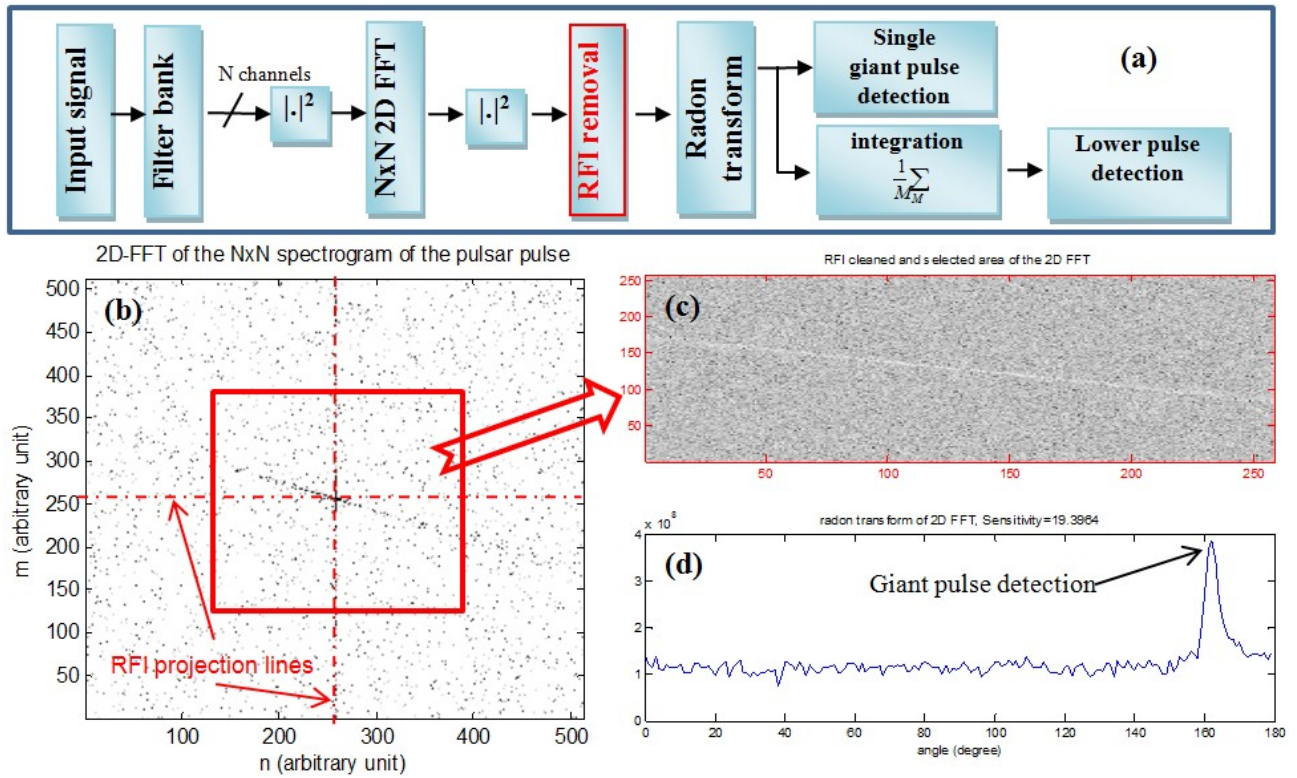


Figure 3: (a) Description de l'algorithme proposé. (b) exemple de la sortie de la FFT 2D 512x512 (l'image a été binarisée pour un meilleure visualisation). les RFIs vont se retrouvés projetés sur la croix rouge centrale. La trainée noire correspond à la signature du pulse géant. Cette signature passe systématiquement par le centre de l'image. Seul l'angle change en fonction du DM. (c) Zone centrale extraite de la FFT 2D après élimination de la croix rouge. Cela permet de maximiser la détection (d) Une transformée de Radon permet de mesurer l'énergie présente sur chaque radiale de l'image 3.c. Un pic d'énergie correspond à la présence d'un pulse géant.

bande étroite vont être concentrés dans les lignes centrales horizontales et verticales. Pour éliminer les RFI, il suffit alors de masquer cette zone particulière avant d'appliquer la suite de l'algorithme. Afin d'augmenter le rapport signal sur bruit, seule la zone centrale de la FFT2D est conservée pour la suite. En pratique, les symétries intrinsèques à la FFT2D sont prises en compte pour limiter la zone traitées.

3. Une transformée de Radon [5] est appliquée sur cette FFT2D nettoyée. Cela consiste à intégrer l'intensité des pixels le long des différentes radiales qui balayent cette image. Si la ligne d'intégration correspond à la ligne où ont été concentrés les pulses du pulsar, alors un pic de détection apparaîtra.

L'avantage de cette approche par rapport à l'approche classique est que tout peut être traité en temps réel. De plus, la méthode présente une certaine robustesse intrinsèque aux RFI et ne nécessite aucune connaissance préalable sur le pulsar. Elle est donc très bien adaptée pour faire de la recherche de nouveaux événements. Dans la section suivante, nous allons analyser la sensibilité de cette approche comparativement à l'approche classique.

3. Calcul des sensibilités

Nous avons calculé les sensibilités respectives de l'approche classique et de l'approche proposée. Cette comparaison est fondée sur un scénario simplifiée (taille du plan temps-fréquence $N \times N$, nombre de plans intégrés M , pulse rectangulaire de longueur L , $DM=0$ (pas de RFI), rapport pulsar sur bruit $SNR < 1$). Les sensibilités théoriques sont alors les suivantes (la définition de la sensibilité est donnée à la figure 4):

pour l'approche classique :
$$S_{classic} = \frac{\sigma_p^2}{\sigma_n^2 / \sqrt{NM}} = SNR \sqrt{NM}$$

pour l'approche proposée :
$$S_{2DFFT} = T \frac{\sigma_p^4}{\sigma_n^4 / \sqrt{NM}} = T \cdot SNR^2 \sqrt{NM}$$

avec T est la largeur du pulse dans le plan temps-fréquence $N \times N$ ($T=L/N$), σ_p^2 est la puissance de pulse, σ_n^2 est la puissance du bruit, le SNR est défini comme σ_p^2 / σ_n^2 . Le pulse et le bruit sont des signaux gaussiens centrés.

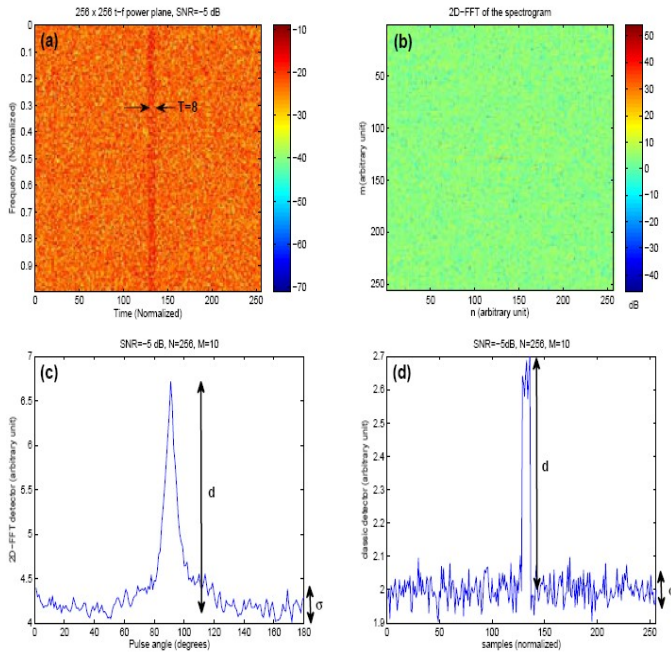


Figure 4: Scénario et définition de la sensibilité pour la comparaison des approches classiques et 2DFFT/Radon. (a) Représentation temps-fréquence du pulse de largeur T avec $DM=0$. Dans cet exemple, $SNR=-5dB$, $L=2048$, $N=256$ et $T=L/N=8$. (b) FFT 2D du plan temps fréquence. Le pulse est projeté sur la ligne horizontale centrale. (c) Profil intégré de $M=10$ transformées de Radon. (d) Profil classique de $M=10$ pulses intégrés. Pour les 2 profils, la sensibilité est définie par le rapport d/σ .

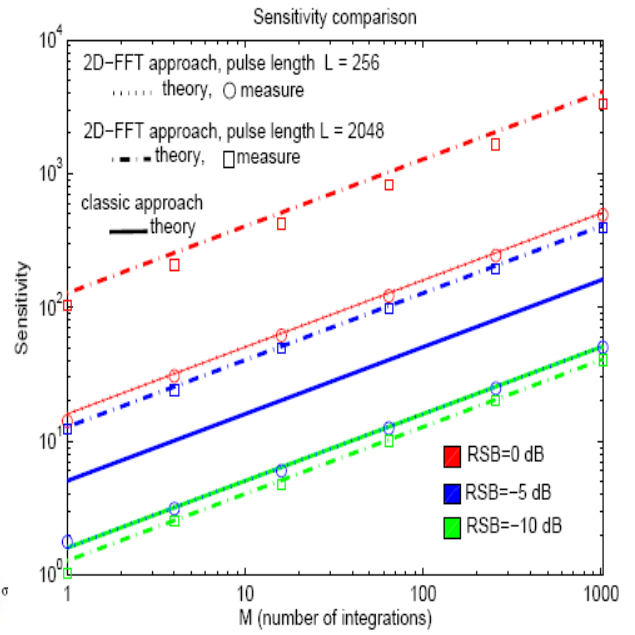


Figure 5: Comparaison des sensibilités pour les 2 procédures ($N=256$). 3 SNR (0 dB, -5dB, -10 dB) et 2 longueurs de pulses ($L=256$ et $L=2048$) ont été simulés pour différentes valeurs de M (nombre de profils intégrés). Les courbes continues et en pointillé correspondent aux résultats théoriques. Le carrés et les ronds correspondent à la mesure des sensibilités de la méthode FFT 2D/Radon par simulation.

A la figure 5, ces sensibilités théoriques ont été comparées à des mesures pratiques obtenues par simulation. Plusieurs remarquent en découlent:

- la théorie modélise correctement la simulation. Le décalage pour le cas « $SNR=0$ dB, $L=2048$ » vient du fait que l'hypothèse $SNR < 1$ n'est plus vérifiée dans cette configuration.
- La sensibilité S_{2DFFT} est proportionnelle au carré du SNR , ce qui rend la méthode moins performante pour les faibles SNR (voir par exemple le cas « $SNR=-5dB$, $L=256$ »). Théoriquement, cette faiblesse peut être compensée par le fait que l'approche proposée prend en compte la longueur T du pulse. Ainsi la performance s'améliore lorsqu'on passe de $L=256$ à $L=2048$.
- La méthode proposée ne suppose connue ni le DM ni la période alors que les performances de la méthode classique reposent expressément sur la connaissance de ces 2 paramètres.

La section suivante décrit l'implantation de cet algorithme de détection en aveugle dans un GPU.

4. Implantation sur GPU

Des tests de performances ont été réalisées pour prouver l'efficacité des GPUs comme une solution appropriée au traitement de données de bande plus large et en temps réel [7].

Concernant la nouvelle méthode de détection en aveugle, elle est en court d'implantation sur une machine de test équipée d'un GPU GeForce GTX 285, ayant 240 cœurs, 1 Go de mémoire globale et 16 Mo de mémoire partagée. Le système d'exploitation utilisé est Linux Gentoo avec la bibliothèque CUDA 2.3.

4.1. L'architecture CUDA

La librairie CUDA est très bien adaptée pour exploiter au mieux les capacités de calculs parallèles du GPU. En effet, quand le GPU est programmé à travers CUDA C, qui est une extension du langage C, il est vu comme un dispositif de calcul, capable d'effectuer un grand nombre d'opération parallèles.

Une portion d'algorithme qui doit être exécutée plusieurs fois indépendamment sur différentes données, peut être isolée et compilée comme un programme à part, nommé Kernel. Une fois chargée dans le GPU, elle sera exécutée N fois par N "threads" ou tâches parallèlement. Les "threads" peuvent être organisés en une, deux ou trois dimensions, et peuvent être synchronisés à l'intérieur d'un bloc qu'on appelle "thread block". Chaque "thread" est identifié par son ID, qui est

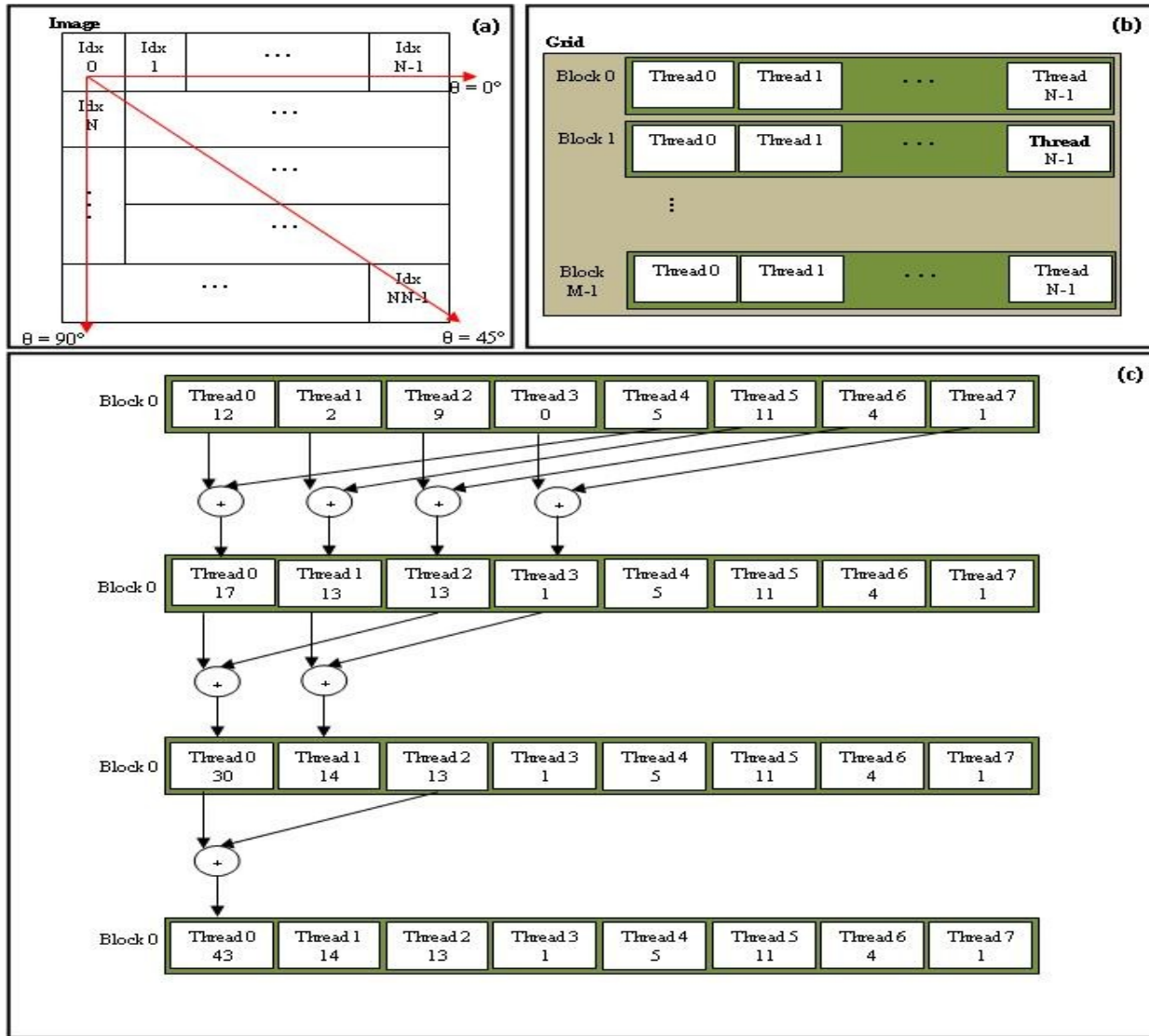


Figure 6 : Implémentation de Radon dans le GPU. (a) : organisation des indices de l'image (sortie de la FFT2D), de gauche à droite et de haut en bas. Les lignes radiales balayent l'image entre 0° et 90° avec un pas donné, au total M valeurs de θ . (b) : Organisation du Grid. Chaque bloc représente une valeur de θ . Chaque thread dans le bloc charge un pixel de l'image. (c) : Schéma expliquant l'algorithme de réduction. Dans chaque bloc, les valeurs des premiers $N/2$ threads s'additionnent avec les derniers $N/2$ en sauvant les résultats dans les $N/2$ threads. Ainsi le nombre de threads actifs est divisé par 2. Ainsi de suite, jusqu'à ce que le nombre de threads actifs soit égal à 1, le résultat final de la somme est récupéré au thread numéro 0.

son numéro dans le bloc. Finalement, les blocs eux même peuvent être organisés de la même manière dans une grille nommée "Grid" [8]. Le nombre de blocs et le nombre de threads doivent être donnés à l'appel du Kernel, tout en respectant certaines règles énoncées dans [8]. La règle la plus importante est que le nombre de registres utilisés par thread multiplié par le nombre de threads dans le bloc ne doit pas dépasser le nombre total des registres par multiprocesseur. Sinon, le bloc ne peut pas être exécuté et le Kernel correspondant ne va pas se lancer.

4.2. Mise en œuvre de Radon

Dans la nouvelle méthode expliquée dans le paragraphe 1, il y a deux étapes clés nécessaires à sa mise en œuvre. La transformée de Fourier à deux dimensions et la transformée de Radon. La FFT2D peut être réalisée avec la bibliothèque CUFFT fournie avec CUDA. Ainsi dans ce paragraphe, on se focalisera sur la mise en œuvre de la transformée de Radon.

En fait, la transformée de Radon qui nous intéresse est simplement la somme des intensités dans l'image suivant des lignes radiales passant par le centre de l'image. Appelons θ l'angle de cette radiale avec l'horizontal. Compte tenu de la symétrie présente dans l'image en sortie de la FFT2D, et pour minimiser le temps de calcul, la transformée de Radon est appliquée uniquement sur le quart de l'image entre 0° et 90° .

Alors, pour chaque valeur de θ , les indices des pixels de la ligne radiale correspondante sont calculés à l'avance. Ensuite, les indices des M angles θ choisis sont stockés dans une matrice qui sera transférée une fois pour toute dans la mémoire partagée du GPU. Cette technique permet d'optimiser les transferts de données et elle améliore la rapidité de traitement.

Concrètement, le nombre de blocs est égal au nombre des valeurs de θ choisies, et le nombre de threads est égal au nombre de pixels pour chaque θ (cf. figures 6.a et 6.b). Ensuite, leur somme est calculée par l'algorithme de réduction expliqué par le schéma de la figure 6.c. Le résultat de la somme des pixels le long d'une ligne radiale est récupéré à partir du thread numéro 0 de chaque bloc. Le résultat final est un vecteur de taille M contenant la valeur de l'intensité cumulée du spectre temps-fréquence sur chaque radiale. La détection d'un pic permet une détection en aveugle d'un pulse dans le plan temps-fréquence comme cela est illustré par la figure 3.d.

5. Conclusion

De part son efficacité en terme de ressource matérielle et de part sa robustesse aux interférences, l'approche FFT2D/Radon est une alternative intéressante à la recherche de nouveaux signaux impulsionnels notamment pour les radiotélescopes de nouvelles générations tels que LOFAR ou SKA où des quantités gigantesques de données devront être traitées. Elle peut servir à dégrossir le travail d'analyse lors de la recherche de nouveaux phénomènes impulsionnels en fournissant en temps réel des détections potentielles.

Remerciements : Les auteurs voudraient remercier la commission européenne du FP7 (projet PrepPSKA (contrat n° 212243) et le projet Radionet UNIBOARD (contrat n° 227290)), l'ANR (contrat n° ANR-09-BLAN-0225-04) pour le financement d'une partie de ce travail.

Références bibliographiques

- [1] D. Lorimer and M. Kramer “*Handbook of pulsar astronomy*,” Cambridge University Press, 2005.
- [2] D. Ait-Allal, C. Dumez-Viou, R. Weber, G. Desvignes, I. Cognard, and G. Theureau, “*RFI mitigation at Nançay Observatory: Impulsive Signal Processing*”, Widefield Science and Technology for the SKA, S. A. Torchinsky, A. van Ardenne, T. van den Brink-Havinga, A. van Es, A.J. Faulkner (eds.), ISBN 978-90-805434-5-4, 2010.
- [3] R. P. Eatough, E. F. Keane, and A. G. Lyne, “*An Interference Removal Technique for Radio Pulsar Searches*,” *Mon. Not. R. Astron. Soc.*, Jan. 2009.
- [4] McLaughlin et al., « *Rotating Radio Transients* » *Nature* 439, 817, 2006.
- [5] S. Deans, “*The radon transform and some of its applications*” Krieger Pub. Co., 1983
- [6] <http://casper.berkeley.edu/wiki/ROACH>
- [7] Thèse de Grégory Desvignes. « L'observation des pulsars au Radiotélescope de Nançay. Applications à la recherche de nouveaux objets, à l'étude des systèmes binaires relativistes et à la détection d'un fond d'ondes gravitationnelles », 2009.
- [8] NVIDIA CUDA Compute Unified Device Architecture, Programming Guide, version 2.0, 2008.